Bayesian variable selection for genetic association in case–control studies after accounting for population stratification

Miao-Yu Tsai (蔡秒玉)
Institute of Statistics and Information Science, College of Science,
National Changhua University of Education

Abstract

Assessing the population-based association between genetic variants and a complex disease can lead to a spurious association caused by population stratification. In order to control the confounding effects of population stratification, we employ generalized linear mixed models (GLMMs) with heterogeneity across populations in case-control genetic studies. For the purpose of Bayesian variable selection, we develop a "higher posterior model probability with bootstrap" (HPMPB) approach to select true genetic covariates and gene–gene (GG) interactions without fitting all possible models. In addition, we derive an efficient approximation approach with Laplace's method and Taylor's expansion to calculate marginal likelihoods to accelerate the computation of the HPMPB selection procedure. Simulation studies are conducted to compare the performance between the HPMPB, the highest posterior model probability (HPMP) and the extended Bayesian model averaging (EBMA) selection approaches, and an application of HapMap data is illustrated. These results reveal that the proposed selection approach is computationally feasible and reliable for exploring potential candidate genes and GG interactions, after adjusting for complex structures among populations.